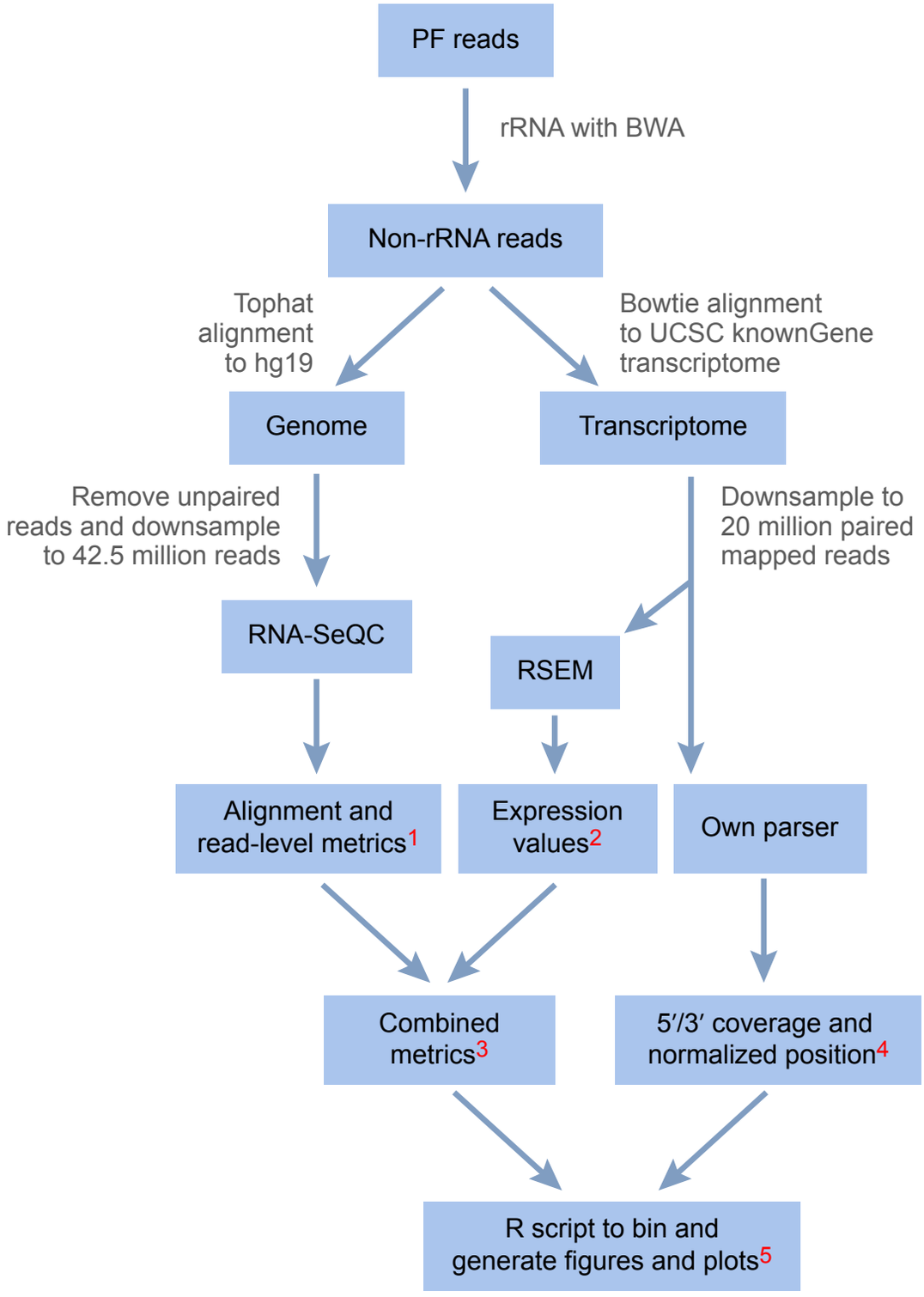
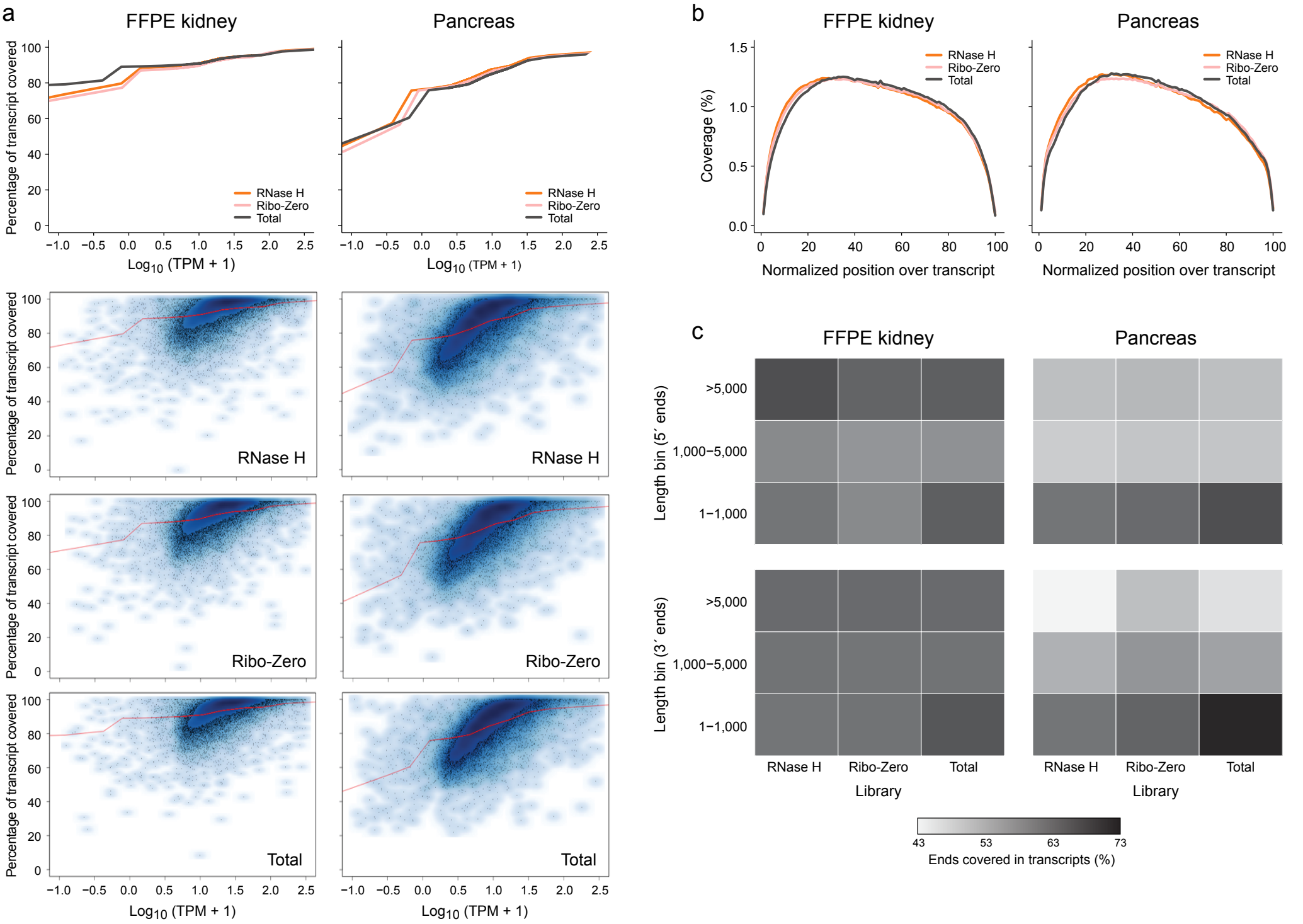
Suppl. Figure 7.
Computational analysis work flow.

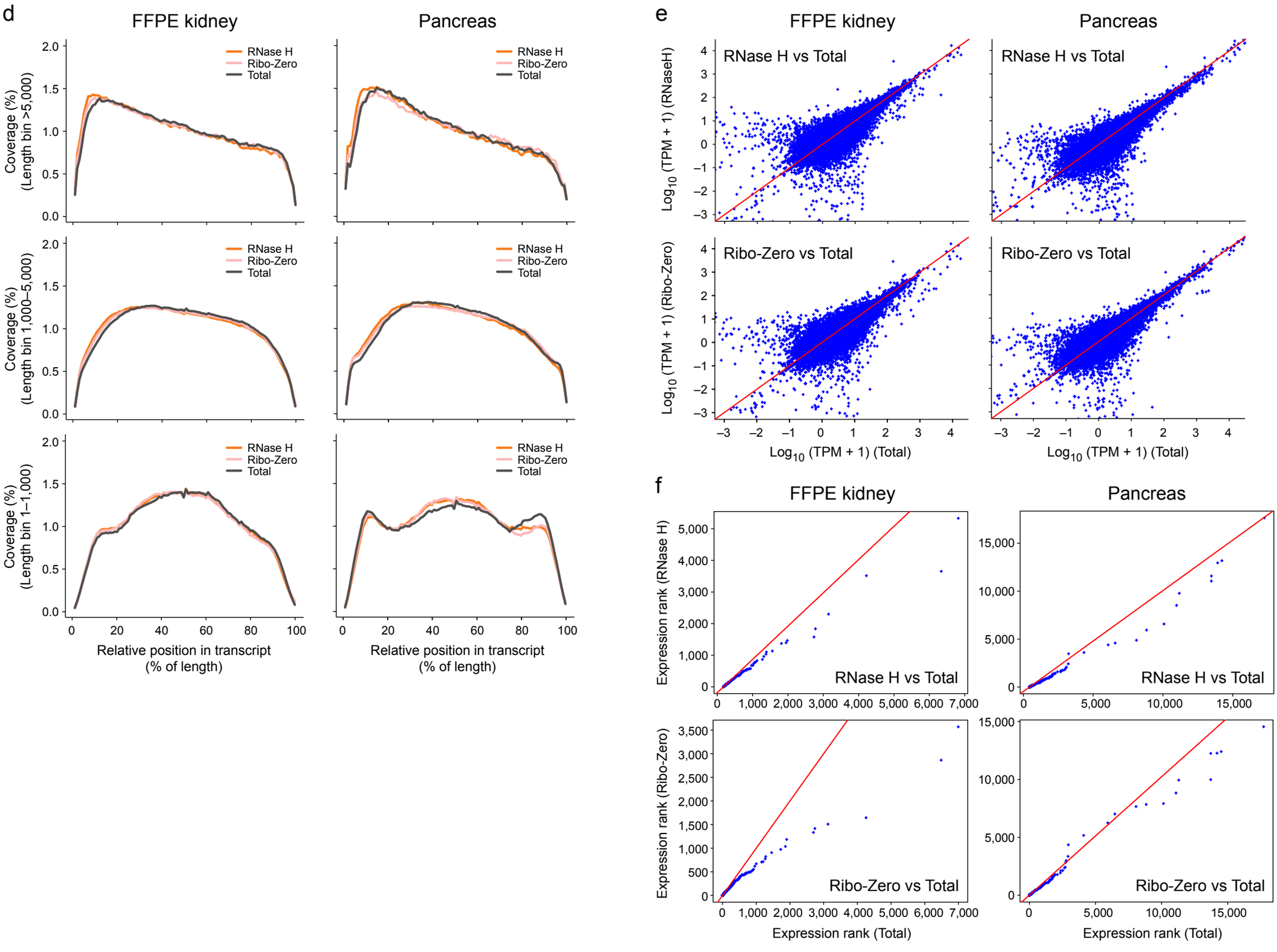


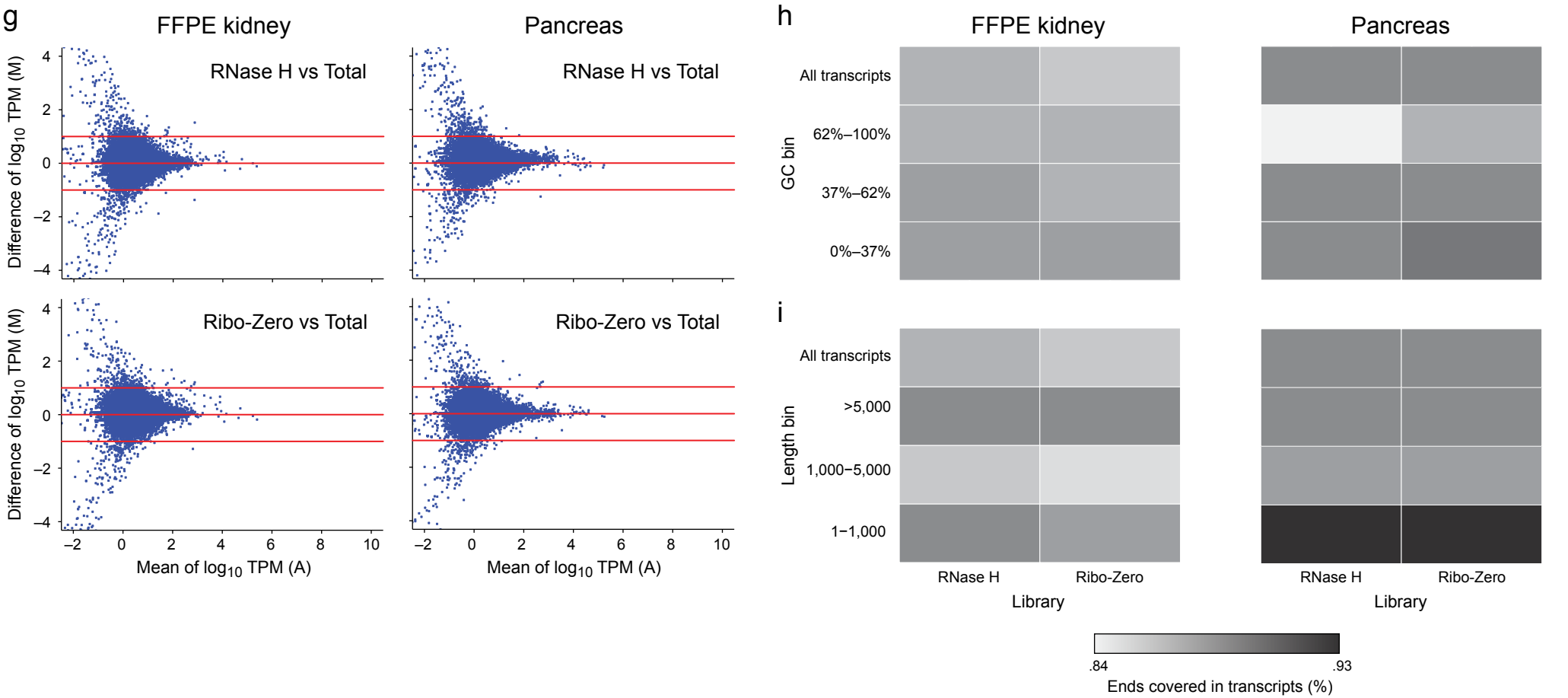
1. RNA-SeQC was used to get the following metrics: number of genes detected, duplication rate, read alignment positions (exonic, intronic, intergenic), coefficient of variation, percent gene covered, and coverage gaps (≥ 5 adjacent bases without coverage).
2. Isoform-corrected expression values. RSEM produced tau values. We multiplied these by 1,000,000, which produes FPKM-like expression estimates.
3. Using *R*, we aggregated the metrics on a per-transcript basis.
4. Each transcript that was expressed in Total library was defined as covered if at least one read mapped to its first or last 100 base pairs. To get a normalized position we binned reads from each transcript with at least 500 reads in the Total library into 100 equally sized bins. We then calculated the relative coverage compared with the coverage for the entire transcript.
5. Figures and plots of all measures mentioned were generated using ggplot in *R*.

Suppl. Figure 8.

Additional performance metrics for actual degraded samples.

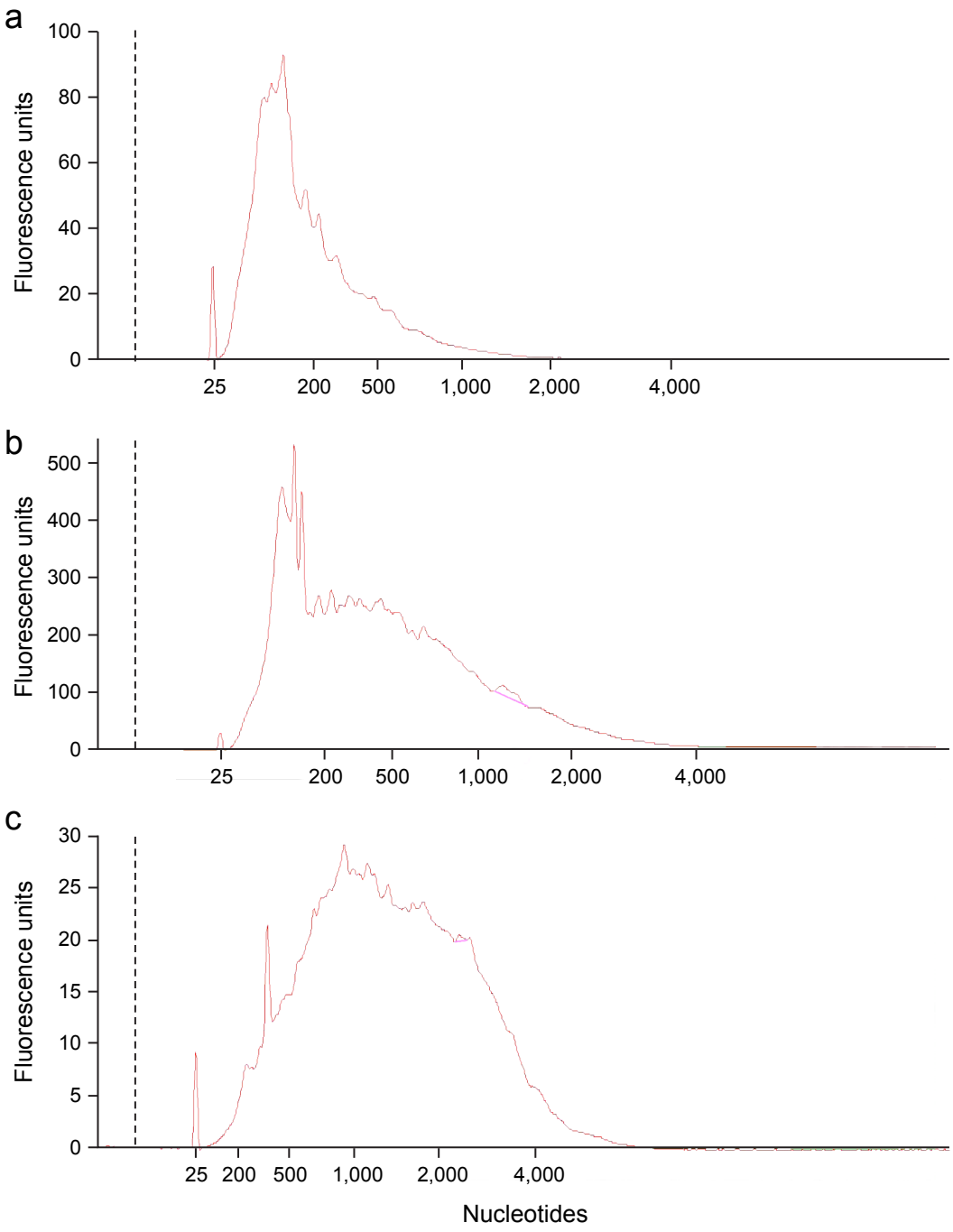



d
FFPE kidney
Pancreas
RNase H
Ribo-Zero
Total
Coverage (%)
(Length bin >5,000)
Coverage (%)
(Length bin 1,000–5,000)
Coverage (%)
(Length bin 1–1,000)
Relative position in transcript
(% of length)
e
FFPE kidney
Pancreas
RNase H vs Total
Ribo-Zero vs Total
Log10 (TPM + 1) (RNaseH)
Log10 (TPM + 1) (Ribo-Zero)
Log10 (TPM + 1) (Total)
f
FFPE kidney
Pancreas
RNase H vs Total
Ribo-Zero vs Total
Expression rank (RNase H)
Expression rank (Ribo-Zero)
Expression rank (Total)



(**a**) Proportion of transcript covered at each expression level. Shown are the Lowess fits of the percentage of the transcript length covered (*y*-axis) for transcripts at each expression level (*x*-axis). Transcript coverage was aggregated for all isoforms of each gene. For each library, shown is the proportion of each transcript covered by reads (*y*-axis, blue dots) at each expression level (*x*-axis), as well as the Lowess fits of this data (red curve). We use a density diagram to indicate the number of genes in each portion of the plot. (**b**) Normalized coverage by position. For each library, shown is the average relative coverage (*y*-axis) at each relative position along the transcripts' length. (**c**) Effect on coverage of 5′ and 3′ ends. Shown are the percent of 5′ (left) and 3′ (right) ends (color scale, far right) in each library (columns) for transcripts with different lengths (rows). (**d**) Effect on normalized coverage by position. For each library, shown is the relative coverage (y-axis) at each relative position along the transcripts' length for short (left), medium (middle) and long (right) transcripts. (**e–g**) Expression level metrics shown as scatter plots (**e**), Q-Q plots (**f**), and MA plots (**g**) between each rRNA-depleted library (*y*-axis) and the control Total library (*x*-axis). (**h,i**) Shown are the Pearson correlation coefficients between each library (columns) and the control Total library for either all transcripts (top row) or for transcripts with (**h**) different lengths; or (**i**) different GC content.

Suppl. Figure 9.
Integrity of degraded RNA samples.



Shown are the BioAnalyzer plots for degraded RNA samples: (**a**) FFPE kidney, (**b**) pancreas, (**c**) fragmented K-562.